

Db	Accession	Organism	Length
Db	AGAGGGCGCTTGTGTCACAGGAGCGCGGCGGATCTTCTCCGGCATGAGAGGACGAGCC	Human	65
QY	gctggtctctccctcaactcctgaagtgctgctctgctctctgcaactgccgagcc	Human	123
Db	GTGCGCTTCCTCCCTCCACCTCCGTAAGGTGCTGCTCCCTGCTGCGACCTGCGGAGCC	Human	125
QY	124 cagatgtgacatcagagcccccactcagagagccctctctctctactcagatcgaagc	Human	183
Db	126 CAGGATTCACACTAGGCTCCACTCCACTCCAGGACCCCTCTCTCTACCGAATACGAAACG	Human	185
QY	184 ttcttcgacatgctgactcactccaactctgaagagcagaatactatgctgctcgttgcaac	Human	243
Db	186 TTCTTGCGACTGCTGACTCCACTCCAGTGAAGGACAGACTACTCTCCCTCCGTCGAACCC	Human	245
QY	244 caagctgctcgggaatcccaactgctgctcagcttggaaccaataatgaaacccagcttagtg	Human	303
Db	246 CAGGCTGCGGGAATCCCACTCTCCAGCTGAGACCAATGTGAACCAACAGGCTTAGTG	Human	305
QY	304 ccgcatggtgctgtctgtctcacaactccctatgactctcgtgttgtagtctctctgacag	Human	363
Db	306 CCGGAT-----	Human	311
QY	364 tctactactacactgtgtgtccacaacagcttactatgctcgaagagatgctgtgtccag	Human	423
Db	312 -----GAGAGTCTGTGTCTCCAG	Human	330
QY	424 ccagctctctatctctcactactaactctcacaagagatagaaagctcagctgaagtcca	Human	483
Db	331 CCAATCTCTATTCTCTACACTTAACACTCTCAAGGAGATAGAAAGCTTCAGCTGAAGTCTCA	Human	390
QY	484 cccacacagatgactctccccaactcacaaccactcacaagtgaacaaagccagaccttc	Human	543
Db	391 CCCACACAGATGACTCTCCCACTCTCACCCTTCACAGTGAACAGAACGCCAAGACCTTC	Human	450
QY	544 cagccctggtcctgaagaggttctcagaacaaacgfggaagagctcctcaaatctctcttcc	Human	603
Db	451 CAGCCCTGGCTTAGAGGCTCAACAACAAGGAGAGGAGCTCTCAATCTCTCTTGTCC	Human	510
QY	604 ctggaagcaggaagagaacgcacagagcaaacagcagagcagaagatggaacacagagcag	Human	663
Db	511 CTGGGAGGCGCAGGAGCAAGGCCAGAGCAACAGCAGAGGAGTGAACACACAGCGAG	Human	570
QY	664 gacccgacacagaacacaagcagaagaggggcagagaaacaggaagcacaagaagagaa	Human	723
Db	571 GAGCGGACACAAACACAAAGCAGGAGGAGGCGGCAAGAAACAGGAAGACCAAGCAGAGAC	Human	630
QY	724 caggaagaggaagaggaagcaggaagaaga-----cagggagcctaaggaagggagcgg	Human	775
Db	631 CAGGACACAGGAGGGAACACGACGAGAAAGAACGACGAGGAGGACCTAAAGAGGAGGAGG	Human	690
QY	776 agagctgtgtctcagctgacagacaga--ctcagaagcccaggttcaactctgaatctatc	Human	833
Db	691 AGGCTGTGTCTCAGCTGACACGACGAACTCAAGAGAGCAAGTTTCACTCTGAATCTTATC	Human	750
QY	834 ttctaaccttccctctgtctcccggttagagaagatagatctactcctatga	Human	889
Db	751 TTCTAACCTCTCTCTTGTCTCCCGGGTTACACAAATAGTCTACTCTATGA	Human	806

REFERENCE	TITLE	JOURNAL	COMMENT
1 (bases 1 to 529)	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .		
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index			
Unpublished (1997)			
Contact: Robert Strausberg, Ph.D.			
Email: <a href="mailto:cgaps-internal.nih.gov">cgaps-internal.nih.gov</a>			
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.			
cDNA Library Arrayed by: Greg Lennon, Ph.D.			
DNA Sequencing by: Washington University Genome Sequencing Center			
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/MLM at: <a href="http://www.bio.linn.gov/bdrip/image/image.html">www.bio.linn.gov/bdrip/image/image.html</a>			
Insert Length: 1076 Std Error: 0.00			
Seq primer: -40ml3 fwd. ET from Amer sham			
High quality sequence stop: 411.			
Location/Qualifiers			
1. 529			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:1641675"			
/clone_lib="Soares_testis_NHT"			
/sex="male"			
/lab_host="DH10B"			
/note="Vector: pT773D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCGCCAATTTTATTTTATTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT773 vector. Library went through one round of normalization to Cot5. and was constructed by Bento Soares and M. Fatima Bonaldo."			
BASE COUNT	80 a	143 c	134 g
ORIGIN			172 t

Query Match	27.9%;	Score 529;	DB 10;	Length 529;
Best Local Similarity	100.0%;	Pred. No. 2.9e-104;		
Matches 529;	Conservative	0;	Mismatches 0;	Indels 0;
			Gaps	0;
QY 454 aaggagatgaagcttcacgtcagtcgaagctcaccacacacagatgacctcccccattccacc				513
DB 529 AAGGAGATGAAGCTTCACTGATAGTCTACCCACACAGATGACCTCCCCATCTACCC				470
QY 514 cacttcacagctacagaaagccagaccttcacgcccctggccttgagaggtcagcaaac				573
DB 469 CACTTCACAGTGCAGAAAGCCAGACCTTCCAGCCCTGGCTGAGAGGCTCAGCAACAAC				410
QY 574 gtggagaagctcctaaatcctcctctgtctcctcgtggagagccagagaaagccagagcac				633
DB 409 GTGGGAAGAGCTCTTAACAATCTCTCTTGTCCCTGGGAGGCCAGAGGCAAGGCCAGACAC				350
QY 634 aagcagagcaaaagagtggagacacagcagagagccgacacaaagaaacacaaacagcaggaag				693
DB 349 AAGCAGAGCAAAAGAGTGGAGCACAAGCAGAGGACCGACACAAGAACACAAGCAGGAAGAG				290
QY 694 gggcagaacaaggaagagcagaagaaggaagcaggaagaggaaggaagcagaagaagga				753
DB 289 GGGCAAGAAACAGGAAGAGCAAGCAAGAGGAACAGGAAGAGGAAAGCAGGAAGAAAGA				230
QY 754 caaggagcaaaagagagcagggagagctgtgtctcgaactgagacagaactagagcccaag				813
DB 229 CAGGGGACTAAGGAGGAGCGGAGGCGTGTCTCAGCTCGACACACACTAGACCCCAAG				170
QY 814 ttccactctgaacctctatcttctaacccttcctcttctgctcccggtgtaagagaagta				873
DB 169 TTTACGCTGAATCTGATCTTAAACCTTCTCTTTTCTCCCGGATCAGGAAGTA				110
QY 874 gagcttaacctatgatataatgagagaacatccagagagctatcttgatcatgagcccaagaaata				933

DB 109 GAGCTACTCTTATGTAATNGAGAAATCCAGAGCTCATTCGATCCAGCAATA 50  
 QY 934 gatgaatgaatgaatataatgatgaactctactctgagaaccaa 982  
 DB 49 GATGAATGATGAATATATATATGATGAGAACTCTACTGAGAAACCAA 1

RESULT 7  
 AA424694 496 bp mRNA EST 16-OCT-1997  
 LOCUS 213110.1 Soares\_testis.NHT Homo sapiens cDNA clone IMAGE:731779  
 DEFINITION 3' similar to TR:G475021 G475021 SP32 PRECURSOR; contains element  
 TARI repetitive element; mRNA sequence.  
 ACCESSION AA424694  
 VERSION AA424694.1 GI:2102744  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 496)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,  
 T., Waterston, R. and Wilson, R.  
 Mashu-Merck EST Project 1997  
 Unpublished (1997)  
 CONTACT: Wilson R.  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: -41m13 fwd. ET from Amersham  
 High quality sequence stop: 313.  
 Location/Qualifiers

## FEATURES

source  
 1. 496  
 /organism="Homo sapiens"  
 /db\_xref="GDB:5928879"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:731779"  
 /clone\_11b="Soares\_testis\_NHT"  
 /sex="male"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 , Inc., and primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCGAAGTGGAGCGGCCCAATTTTCTTTT 3']  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization to Cos5, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 75 a 134 c 127 g 160 t  
 ORIGIN

Query Match 26.0%; Score 492.8; DB 10; Length 496;  
 Best Local Similarity 99.6%; Pred. No. 1.9e-96;  
 Matches 494; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 487 accaagatgaactcccatcaccatcaccatcagaagacagccagacattcag 546  
 DB 496 ACCAGATGCTCCCTCCATCTCACCCCTTCACAGAGAGAGCCAGACCTTCAG 437  
 547 ccctgagcctgagagctgagcaaacgtggaagagcctcctcctctgctcctg 606  
 36 CCCTGCTGAGAGGCTCAGCAACAGTGGAGAGCTCTTACATCTCTCTGCTG 377

QY 607 ggaagccagagagcagagccagagcacaaagcagagagagtgagcagcagag 666  
 DB 376 GGAAGCCAGAGGAGCAAGCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 317  
 QY 667 ccgacacaaagacacaaagcagagagagagcagaaagcagaaagcagaaagcag 726  
 DB 316 CCGACACAAAGACAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 257  
 QY 727 gaagagagagagagcagagagagagagagagagagagagagagagagagag 786  
 DB 256 GAAAG 197  
 QY 787 cagctgcag 846  
 DB 196 CAGCTGCAGACAGACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 137  
 QY 847 tctttgtcccgagtagcagagagtagagtagctactcctatgataatgagagag 906  
 DB 136 TCTTTTGTCTCCCGGAGTACAGAGAGTACAGTACTCTATGATGAGAGAGAG 77  
 QY 907 gagctcattcagatcagccagagagagagagagagagagagagagagagagag 966  
 DB 76 GAGCTCATTCATCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 17  
 QY 967 tactgag 982  
 DB 16 TACTGAGAGAAACCAA 1

RESULT 8  
 AA824340 496 bp mRNA EST 31-DEC-1998  
 LOCUS a30e09.s1 Soares\_testis\_NHT Homo sapiens cDNA clone 1391848 3'  
 DEFINITION similar to FR:Q60485 Q60485 SP32 PRECURSOR; mRNA sequence.  
 ACCESSION AA824340  
 VERSION AA824340.1 GI:2896410  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 496)  
 NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cga@bcrfemail.nih.gov  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
 , Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CCAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www.bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 609 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham.  
 Location/Qualifiers

## FEATURES

source  
 1. 496  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="1391848"  
 /clone\_11b="Soares\_testis\_NHT"  
 /sex="male"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 , Inc., and primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCGAAGTGGAGCGGCCCAATTTTCTTTT 3']  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library

to identify individuals or in diagnostic procedures to identify individuals having genetic diseases resulting from abnormal expression of the genes corresponding to the extended cDNAs. They are also useful for constructing a high resolution map of the human chromosomes. They can also be used for gene therapy to control or treat genetic diseases.

Sequence 964 BP; 258 A; 294 C; 227 G; 180 T; 5 other:

Query Match 49.4%; Score 937; DB 20; Length 964;  
Best Local Similarity 99.3%; Pred. No. 6,3e-235;  
Matches 948; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

OY 30 gggcgagatttccggcattgagaaagcagccgctgctccctccatccatcgaa 89
    |||||||
DB 2 gggcgagatttccggcattgagaaagcagccgctgctccctccatccatcgaa 61
OY 90 ggtgcgtctctgctctgctgacacgctgacgagccagattcgactcagcccatcc 149
    |||||||
DB 62 ggtgcgtctctgctctgctgacacgctgacgagccagattcgactcagcccatcc 121
OY 150 aggcagccctctctctccacgaataagaaagctcttcgactgctgactcgaactg 209
    |||||||
DB 122 aggcagccctctctctccacgaataagaaagctcttcgactgctgactcgaactg 181
OY 210 gaagcagagactactgctgctgctgacacgagctgacgagaaatccacatcgt 269
    |||||||
DB 182 gaagcagagactactgctgctgctgacacgagctgacgagaaatccacatcgt 241
OY 270 ccagctgagcaaatatgaaacacacgagctgacgagaaatgctgctgctcgaact 329
    |||||||
DB 242 ccagctgagcaaatatgaaacacacgagctgacgagaaatgctgctgctcgaact 301
OY 330 ccctatgctctctgctgctgctgacacgagctgacacacacgctgctgctcgaact 389
    |||||||
DB 302 ccctatgctctctgctgctgctgacacgagctgacacacacgctgctgctcgaact 361
OY 390 cgtctactatgcaagagagctgctgctgctcgaacgagctgctatctcactcaaac 449
    |||||||
DB 362 cgtctactatgcaagagagctgctgctgctcgaacgagctgctatctcactcaaac 421
OY 450 tctcaagagatagaagctcagctgaagctcagccacac-gatgactcccatc 508
    |||||||
DB 422 tctcaagagatagaagctcagctgaagctcagccacacgagactcccatc 481
OY 509 caccacactcagctgacgagacacgacacacacacacacacacacacacacacac 568
    |||||||
DB 482 caccacactcagctgacgagacacgacacacacacacacacacacacacacacac 541
OY 659 acaacgtgagagagctcctcactcctctgctcctgctgagagcagcagcagcag 628
    |||||||
DB 542 acaacgtgagagagctcctcactcctcctcctcctcctcctcctcctcctcctcctc 601
OY 659 agcacaagcagagcagagagctgagcagcagcagcagcagcagcagcagcagcag 688
    |||||||
DB 602 agcacaagcagagcagagagctgagcagcagcagcagcagcagcagcagcagcag 661
OY 669 aagagagcagagcagagcagagcagagcagagcagagcagagcagagcagagcag 748
    |||||||
DB 662 aagagagcagagcagagcagagcagagcagagcagagcagagcagagcagagcag 721
OY 749 aagagcagagcagagcagagcagagcagagcagagcagagcagagcagagcagag 808
    |||||||
DB 722 aagagcagagcagagcagagcagagcagagcagagcagagcagagcagagcagag 781
OY 809 ccaagttcactctgaatctatctctcactcctcctcttgcctcccgagtcagag 868
    |||||||
DB 782 ccaagttcactctgaatctatctcactcctcctcttgcctcccgagtcagag 841
OY 869 aagtgagctactctctctatgataagagagacatccagagctatccatccacagag 928
    |||||||
DB 842 aagtgagctactctctctatgataagagagacatccagagctatccatccacagag 901

```

OY 929 aatagatgaatgaatgaatataatgagaaactcctactgagaaaccaa 983  
DB 902 aatagatgaatgaatgaatataatgagaaactcctactgagaaaccaa 956

## RESULT 7

AAV88163 standard; cDNA: 308 BP.

AAV88163;

12-FEB-1999 (first entry)

EST clone GA63.

Expressed sequence tag; secreted protein; haematopoiesis regulator;  
tissue growth; actinin; inhibitor; tumour invasion suppressor; EST; human;  
chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

Homo sapiens.

WO9845437-A2.

15-OCT-1998.

10-APR-1998; 98WO-US06956.

10-APR-1997; 97US-0837312.

(GENE) GENETICS INST INC.

Agostino M, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

Racie LA, Spaulding V, Treacy M;

MP1: 1999-070078/06.

New polynucleotides encoding human secreted proteins - derived from  
e.g. human blood, kidney, foetal lung, placenta, testes, brain,

ovary, placenta, retina and colon cDNA libraries

Claim 1; Page 304; 641pp; English.

The present sequence represents an expressed sequence tag (EST), and is  
a polynucleotide of the invention. The polynucleotides of the invention  
are all secreted EST sequences and proteins encoded by them are predicted to  
have useful biological activities which would make them suitable for  
treating, preventing or ameliorating medical conditions in humans and  
animals, although no supporting data is given. Suggested activities  
include nutritional activity, immune stimulating or suppressing activity,  
haematopoiesis regulating activity, tissue growth activity,  
activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
activity. The EST sequences are also stated to be useful for gene  
therapy.

Sequence 308 BP; 85 A; 100 C; 75 G; 48 T; 0 other:

Query Match 14.9%; Score 282.4; DB 20; Length 308;  
Best Local Similarity 96.3%; Pred. No. 3.8e-64;  
Matches 289; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

OY 335 actatgcaagagagctctgtgctcagcagcagctctatctcctcaactcaactca 454
    |||||||
DB 3 atcgccctcattagcgcagagctgctcagcagcagctctatctcctcaactcaactca 62
OY 455 aggaatagaagctcagctgaagctcaccacacacagatgacctcccatcacc 514
    |||||||
DB 63 aggaatagaagctcagctgaagctcaccacacacagatgacctcccatcacc 122

```

